NONLINEAR SPECTRAL UNMIXING USING RESIDUAL COMPONENT ANALYSIS AND A GAMMA MARKOV RANDOM FIELD

Yoann Altmann[(1)]*, *Marcelo Pereyra*[(2)]†, *Steve McLaughlin*[(1)]‡

[(1)]Heriot-Watt University, School of Engineering and Physical Sciences, Edinburgh, United Kingdom.
[(2)] University of Bristol, School of Mathematics, Bristol, United Kingdom.

ABSTRACT

This paper presents a new Bayesian nonlinear unmixing model for hyperspectral images. The proposed model represents pixel reflectances as linear mixtures of endmembers, corrupted by an additional combination of nonlinear terms (with respect to the endmembers) and additive Gaussian noise. A central contribution of this work is to use a Gamma Markov random field to capture the spatial structure and correlations of the nonlinear terms, and by doing so to improve significantly estimation performance. In order to perform hyperspectral image unmixing, the Gamma Markov random field is embedded in a hierarchical Bayesian model representing the image observation process and prior knowledge, followed by inference with a Markov chain Monte Carlo algorithm that jointly estimates the model parameters of interest and marginalises latent variables. Simulations conducted with synthetic and real data show the accuracy of the proposed SU and nonlinearity estimation strategy for the analysis of hyperspectral images.

***Index Terms*—** Hyperspectral imagery, nonlinear spectral unmixing, residual component analysis, Gamma Markov random field, Bayesian estimation.

1. INTRODUCTION

Spectral unmixing (SU) is one of the most important problems in the analysis of hyperspectral images. This source separation problem consists of recovering the spectral signatures (endmembers) of the materials and their proportions in each pixel of the observed image. The SU problem has been widely studied for applications where pixel reflectances are linear combinations of pure component spectra [1]. However, the linear mixing model (LMM) can be inappropriate for some hyperspectral images, such as those containing sand-like materials or relief. Nonlinear mixing models provide an interesting alternative to overcoming the inherent limitations of the LMM. Most of the models proposed in the hyperspectral image literature can be divided into two main classes [2]. The first class of nonlinear models consists of physical models based on the nature of the environment (e.g., intimate mixture and polynomial [2] models which have been proposed to handle short and long range multiple light scattering effects, respectively). The second class contains more flexible models allowing different kinds of nonlinearities to be approximated. These flexible models can be constructed from neural networks, kernel functions, or post-nonlinear transformations. We refer the reader to [3] for a recent review of nonlinear hyperspectral SU methods.

While the consideration of nonlinear effects can be relevant in specific areas, the LMM is often sufficient to approximate the actual mixing models in some image pixels or homogeneous regions. Consequently, it makes sense to consider nonlinear models that generalize the LMM when using a single model for SU. In [4] an algorithm based on a nonlinear mixing model inspired from residual component analysis (RCA) [5] has been investigated for joint nonlinear SU and nonlinearity detection. In the context of SU of hyperspectral images, the nonlinear effects were modeled by additive perturbation terms characterized by Gaussian processes (GPs). This allowed the nonlinear terms to be marginalized, yielding a flexible model depending only on the nonlinearity energies. In addition, the proposed algorithm partitioned the pixels into homogeneous regions in which the nonlinearities shared the same GP. However, an important limitation of the method proposed in [4] is that it requires specifying the number of classes (nonlinearity levels) present in the image, which is often difficult to determine a priori and very computationally expensive to infer from data (for instance by using information criteria or reversible-jump Markov chain Monte Carlo methods [6]). Moreover, even with an appropriate number of classes, the method [4] produces spatially piecewise constant estimation results that may represent poorly the true nonlinearities, in particular if these exhibit local fluctuations or vary smoothly across some image pixels. To address these limitations, this work proposes to extend the Bayesian model considered in [4] in order to allow the nonlinearity levels to change freely across the scene, while simultaneously ensuring positive nonlinear terms and promoting regularity and smoothness by exploiting spatial correlations.

The remainder of the paper is organized as follows: Section 2 provides the problem statement and recalls the RCA [5] model for hyperspectral image analysis. Section 3 introduces the proposed hierarchical Bayesian model by specifying the likelihood and the proposed prior and hyper-prior distributions for the parameters of the RCA model. Following on from this, Section 4 describes a Markov chain Monte Carlo (MCMC) algorithm for performing Bayesian inference with the proposed model (e.g. compute minimum mean square error estimators for the parameters of interest, as well as measures of uncertainty). Simulation results conducted on real data are shown and discussed in Section 5, and conclusions and perspectives for future work are finally reported in Section 6.

2. PROBLEM FORMULATION

Let $\mathbf{y}_{i,j} \in \mathbb{R}^L$ be the pixel at location (i, j) of an hyperspectral image $\mathbf{Y}$ of size $N_{\text{row}} \times N_{\text{col}}$ observed at L spectral bands. Following an RCA approach [5, 4], we model each image pixel as a

*Part of this work was supported by the Direction Générale de l'Armement, French Ministry of Defence.

†Part of this work was supported by the SuSTaIN program - EPSRC grant EP/D063485/1 - at the Department of Mathematics, University of Bristol. MP holds a Marie-Curie Intra-European Fellowship for Career Development.

‡Part of this work was supported by the EPSRC via grant EP/J015180/1.

linear combination of R known spectra or endmembers $\mathbf{m}_r$, plus an additive perturbation $\boldsymbol{\phi}_{i,j}$ embedding nonlinearities and additive Gaussian noise

$$\mathbf{y}_{i,j} = \sum_{r=1}^{R} a_{r,i,j}\mathbf{m}_r + \boldsymbol{\phi}_{i,j} + \mathbf{e}_{i,j} = \mathbf{M}\boldsymbol{a}_{i,j} + \boldsymbol{\phi}_{i,j} + \mathbf{e}_{i,j}, \quad (1)$$

for all (i,j), where $\mathbf{m}_r = [m_{r,1}, \ldots, m_{r,L}]^T$ is the spectral response of the rth material present in the scene, $a_{r,i,j}$ is its abundance within pixel (i,j) and $\mathbf{e}_n \sim \mathcal{N}(\mathbf{0}_L, \boldsymbol{\Sigma}_0)$ is Gaussian noise with diagonal covariance matrix $\boldsymbol{\Sigma}_0 = \text{diag}\left(\boldsymbol{\sigma}^2\right)$ with elements $\boldsymbol{\sigma}^2 = [\sigma_1^2, \ldots, \sigma_L^2]^T$ (note that matrix and vector notations $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ and $\boldsymbol{a}_{i,j} = [a_{1,i,j}, \ldots, a_{R,i,j}]^T$ have been used on the r.h.s. of (1)). For the nonlinear effects we use the model

$$\boldsymbol{\phi}_{i,j} = \sum_{k=1}^{R-1} \sum_{k'=k+1}^{R} \gamma_{i,j}^{(k,k')} \sqrt{2}\mathbf{m}_k \odot \mathbf{m}_{k'} + \sum_{k=1}^{R} \gamma_{i,j}^{(k)} \mathbf{m}_k \odot \mathbf{m}_k. \quad (2)$$

that is motivated by the fact that nonlinearities in SU can often be represented as polynomial functions of the endmembers [7, 8]. In this paper we do not limit the values of the constraints $\left\{\gamma_{i,j}^{(k,k')}\right\}, \left\{\gamma_{i,j}^{(k)}\right\}$ and focus on estimating *nonlinearity levels*. Also note that due to physical considerations we model the abundances as non-negative quantities $a_{r,i,j} \geq 0$. Because we consider nonlinear SU we do not use the sum-to-one constraint for the abundances that is frequently enforced in linear SU, though this constraint could be easily included in the model described below.

This papers considers the problem of jointly estimating the abundance vectors and the level of nonlinearity in each pixel (represented here by the parameters $\left\{\gamma_{i,j}^{(k,k')}\right\}, \left\{\gamma_{i,j}^{(k)}\right\}$ gathered in the matrix $\boldsymbol{\Gamma}$). This problem is addressed in the Bayesian framework by defining an appropriate statistical model and inference algorithm presented in Sections III and IV below.

3. BAYESIAN MODEL

This section presents an original hierarchical Bayesian model for the unknown quantities of interest; that is, the abundance vectors and the nonlinearity coefficients. For completeness we assume that the noise covariance $\boldsymbol{\sigma}^2$ is also unknown and to be estimated. This Bayesian model is defined by specifying the likelihood and the prior and hyper-prior distribution for the model parameters in (1)-(2).

3.1. Likelihood

From eq. (1) and by assuming that the observations gathered in $\mathbf{Y}$ are conditionally independent we obtain that

$$\mathbf{Y}|\mathbf{M}, \mathbf{A}, \boldsymbol{\Phi}, \boldsymbol{\sigma}^2 \sim \prod_{i,j} \mathcal{N}(\mathbf{x}_{i,j}, \boldsymbol{\Sigma}_0) \quad (3)$$

where $\boldsymbol{\Phi} = \left\{\boldsymbol{\phi}_{i,j}\right\}_{i,j}$ denotes the nonlinear terms in (2), $\mathbf{A}$ is an $R \times N_{\text{row}} \times N_{\text{col}}$ array of abundances and $\mathbf{x}_{i,j} = \mathbf{M}\boldsymbol{a}_{i,j} + \boldsymbol{\phi}_{i,j}$.

3.2. Prior for the abundances A

We assign the abundances the hierarchical prior distribution

$$\begin{aligned} a_{r,i,j}|\beta_r &\sim \mathcal{N}_{\mathbb{R}^+}(0, \beta_r) \\ \beta_r &\sim \mathcal{IG}(\alpha_1, \alpha_2) \end{aligned} \quad (4)$$

parametrised by some fixed hyper-parameters α_1 and α_2, and where we note that the prior on $a_{r,i,j}|\beta_r$ is truncated to $\mathbb{R}^+$ to reflect the positivity of $a_{r,i,j}$. This prior is very flexible and can be adjusted to represent a wide variety of prior beliefs. Here we set $\alpha_1 = 1$ and $\alpha_2 = 2$, leading to a (marginal) exponential prior for $a_{r,i,j}$ modelling that abundances are sparse and often take values in $[0, 1]$. Moreover, we expect the abundances associated with the same material to exhibit correlations, in particular in terms of their scale. This belief is encoded in (4) by defining one hidden variable β_r for each material or endmember $\mathbf{m}_r$, which is shared by all the abundances related to that material. This hierarchical structure operates as a global pooling mechanism that shares information across the abundances associated to each material). Finally, assuming that abundances are prior independent given the hidden variables $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_R]^T$, we obtain $f(\mathbf{A}, \boldsymbol{\beta}) = f(\mathbf{A}|\boldsymbol{\beta})f(\boldsymbol{\beta})$ with $f(\mathbf{A}|\boldsymbol{\beta}) = \prod_{r,i,j} f(a_{r,i,j}|\beta_r)$ and $f(\boldsymbol{\beta}) = \prod_r f(\beta_r|\alpha_1, \alpha_2)$.

3.3. Priors for the nonlinearity coefficients Γ

A central contribution of this paper is to propose the following hierarchical prior for the nonlinearity coefficients

$$\begin{cases} \boldsymbol{\gamma}_{i,j}|s_{i,j} \sim \mathcal{N}_{(\mathbb{R}^+)^K}(\mathbf{0}, s_{i,j}\mathbf{I}_K) \\ s_{i,j} \sim \mathcal{IG}(\alpha_3, \alpha_3\alpha_{4,i,j}), \end{cases} \quad (5)$$

which is parametrized by a local hyper-parameter $\alpha_{4,i,j}$ related to the prior mean of $s_{i,j}$, and therefore to the strength of the nonlinearities at the pixel (i,j), and by a global hyper-parameter α_3 that controls the deviation of $s_{i,j}$ from $\alpha_{4,i,j}$. We wish to specify (5) to reflect that we expect the values of $\boldsymbol{\gamma}_{i,j}$ to vary smoothly from one pixel to another, with occasional abrupt changes. In order to model this behaviour we specify $\alpha_{4,i,j}$ such that the resulting prior for $\boldsymbol{\Gamma}$ is a hidden gamma-Markov random field (GMRF) [9].

More precisely, we denote by $\mathbf{S}$ the $N_{\text{row}} \times N_{\text{col}}$ matrix with elements $s_{i,j}$, introduce a $(N_{\text{row}} + 1) \times (N_{\text{col}} + 1)$ auxiliary matrix $\mathbf{W}$ with elements $w_{i,j} \in \mathbb{R}^+$ and define a bipartite conditional independence graph between $\mathbf{S}$ and $\mathbf{W}$ such that each $s_{i,j}$ is connected to four neighbour elements of $\mathbf{W}$ and vice-versa. With this 1st order neighbourhood structure, any given $s_{i,j}$ and $s_{i+1,j}$ are 2nd order neighbours via $w_{i,j+1}$ and $w_{i+1,j+1}$. We specify a GMRF prior for $\mathbf{S}, \mathbf{W}$ [9], and obtain the following hierarchical prior for $\boldsymbol{\Gamma}, \mathbf{S}, \mathbf{W}$

$$\begin{aligned} \boldsymbol{\gamma}_{i,j}|s_{i,j} &\sim f(\boldsymbol{\gamma}_{i,j}|s_{i,j}) && \text{(6a)} \\ s_{i,j}|\mathbf{W}, \alpha_3 &\sim \mathcal{IG}(\alpha_3, \alpha_3\alpha_{4,i,j}(\mathbf{W})) && \text{(6b)} \\ w_{i,j}|\mathbf{S}, \alpha_3 &\sim \mathcal{G}(\alpha_3, 1/(\alpha_3\alpha_{5,i,j}(\mathbf{S}))) && \text{(6c)} \end{aligned}$$

where

$$\begin{aligned} \alpha_{4,i,j}(\mathbf{W}) &= w_{i,j} + w_{i+1,j} + w_{i,j+1} + w_{i+1,j+1}/4 \\ \alpha_{5,i,j}(\mathbf{W}) &= (s_{i,j}^{-1} + s_{i-1,j}^{-1} + s_{i,j-1}^{-1} + s_{i-1,j-1}^{-1})/4. \end{aligned}$$

The density for this joint prior for $\boldsymbol{\Gamma}$, $\mathbf{S}$ and $\mathbf{W}$ is given by $f(\boldsymbol{\Gamma}, \mathbf{S}, \mathbf{W}|\alpha_3) = f(\boldsymbol{\Gamma}|\mathbf{S})f(\mathbf{S}, \mathbf{W}|\alpha_3)$ where $f(\boldsymbol{\Gamma}|\mathbf{S}) = \prod_{i,j} f(\boldsymbol{\gamma}_{i,j}|s_{i,j})$ and

$$\begin{aligned} f(\mathbf{S}, \mathbf{W}|\alpha_3) \propto & \left[\prod_{(i',j')\in\mathcal{V}_{\mathbf{W}}} w_{i',j'}^{(\alpha_3-1)}\right] \prod_{(i,j)\in\mathcal{V}_{\mathbf{S}}} (s_{i,j})^{-(\alpha_3+1)} \\ & \times \prod_{((i,j),(i',j'))\in\mathcal{E}} \exp\left(\frac{-\alpha_3 w_{i',j'}}{4s_{i,j}}\right). \end{aligned} \quad (7)$$

Notice that we denote explicitly the dependence on the value of α_3, which acts here as a regularisation parameter that controls the amount of spatial smoothness enforced by the GMRF. Following an empirical Bayesian approach, the value of α_3 remains unspecified and will be adjusted automatically during the inference procedure

by maximum marginal likelihood (MML) estimation using the technique [10].

Finally, it is worth mentioning that this model for the nonlinearity coefficients has similarities with the model proposed in [4] that also considers the spatial regularity of non-linearities. However, the model described [4] follows a segmentation approach in which the non-linearity coefficients are assumed (and constrained) to take values in a finite set. This leads to a piece-wise constant representation and requires specification of the number of nonlinearity levels present in the image, a value that is often difficult to determine a priori. The model proposed in this paper provides a spatially smooth representation of the nonlinearities that is possibly more realistic than the piece-wise constant representation of [4], and also has the practical advantage of not requiring practitioners to specify the finite number of admissible nonlinearity levels. Another important distinction is that the model described in [4] does not take into account the positivity of $\boldsymbol{\gamma}_{i,j}$, which we have found to greatly improve the estimation of the nonlinearities.

3.4. Prior for the noise covariance $\boldsymbol{\sigma}^2$

We assume that there is no prior knowledge available about the values of noise covariance (other than the fact that it is diagonal) and assign each diagonal element σ_ℓ^2 a Jeffreys' prior $f(\boldsymbol{\sigma}^2) = \prod_{\ell=1}^{L} f(\sigma_\ell^2)$, with $f(\sigma_\ell^2) \propto \sigma_\ell^{-2} \mathbf{1}_{\mathbb{R}^+}\left(\sigma_\ell^2\right)$.

3.5. Posterior distribution

We are now ready to specify the posterior distribution for $\mathbf{A}, \boldsymbol{\Gamma}, \boldsymbol{\sigma}^2, \mathbf{S}$, $\mathbf{W}$ and $\boldsymbol{\beta}$ given the observed hyper-spectral image $\mathbf{Y}$ and the value of the spatial regularisation hyper-parameter α_3 (recall that this value will be determined by MML estimation during the inference procedure). Using Bayes' theorem, and assuming prior independence between $(\mathbf{A}, \boldsymbol{\beta})$, $(\boldsymbol{\Gamma}, \mathbf{S}, \mathbf{W})$ and $\boldsymbol{\sigma}^2$, the joint posterior distribution associated with the proposed Bayesian model is given by $f(\boldsymbol{\theta}, \boldsymbol{\Psi}|\mathbf{Y}, \mathbf{M}, a) \propto f(\mathbf{Y}|\mathbf{M}, \boldsymbol{\theta}) f(\boldsymbol{\theta}|\boldsymbol{\Psi}) f(\boldsymbol{\Psi}|a)$, with $\boldsymbol{\theta} = (\mathbf{A}, \boldsymbol{\Gamma}, \boldsymbol{\sigma}^2)$, $\boldsymbol{\Psi} = (\mathbf{S}, \mathbf{W}, \boldsymbol{\beta})$), $f(\boldsymbol{\theta}|\boldsymbol{\Psi}) = f(\mathbf{A}|\boldsymbol{\beta}) f(\boldsymbol{\sigma}^2)$ $\times f(\boldsymbol{\Gamma}|\mathbf{S})$, and $f(\boldsymbol{\Psi}|a) = f(\boldsymbol{\beta}) f(\mathbf{S}, \mathbf{W}|a)$.

4. BAYESIAN INFERENCE USING A GIBBS SAMPLER

This section describes an MCMC method to estimate the unknown parameters of the Bayesian model specified in Section III. More precisely, we consider a Gibbs sampler to generate samples distributed according to $f(\boldsymbol{\theta}, \boldsymbol{\Psi}|\mathbf{Y}, \mathbf{M}, a)$; these samples are then used to approximate Bayesian estimators by Monte Carlo integration [11, Chap. 10] (we generate N_{MC} samples for each unknown parameter, remove the first N_{bi} samples associated with the burn-in period, and use the remaining samples to approximate minimum mean square error (MMSE) estimators). The principle of this algorithm is to simulate from the conditional distributions associated with $f(\boldsymbol{\theta}, \boldsymbol{\Psi}|\mathbf{Y}, \mathbf{M}, a)$; we sample simultaneously the abundance and nonlinearity coefficient matrices $(\mathbf{A}, \boldsymbol{\Gamma})$, then the noise variances $\boldsymbol{\sigma}^2$, the nonlineality levels in $\mathbf{S}$, the auxiliary variables $\mathbf{W}$ and finally the hyperparameters $\boldsymbol{\beta}$ using the following 5 successive MCMC steps:

Abundances and nonlinearity coefficients: The conditional distribution of $\boldsymbol{a}_{i,j}, \boldsymbol{\gamma}_{i,j}|\mathbf{y}_n, \mathbf{M}, s_{i,j}^2, \boldsymbol{\sigma}^2$ is a multivariate Gaussian distribution restricted to the positive orthant for the abundances.

Noise variances σ_ℓ^2: Sampling σ_ℓ^2 simply reduces to sampling from an inverse-Gamma distribution. Note that the noise variances are a posteriori independent and can thus be updated in a parallel manner.

Nonlinearity levels: Due to the structure of the GMRF, the nonlinearity levels in $\mathbf{S}$ are (conditioned on all the other variables) posterior independent and can be updated in parallel by sampling univariate inverse-Gamma distributions.

Auxiliary variables: In a similar fashion to the nonlinearity levels, the auxiliary variables $\mathbf{W}$ are (conditioned on all the other variables) posterior independent and can be updated in parallel by sampling univariate Gamma distributions.

Abundance hyperparameters: The hyperparameter vector $\boldsymbol{\beta}$ is also update in parallel by sampling from conjugate inverse-Gamma conditional distributions.

5. SIMULATIONS: REAL HYPERSPECTRAL IMAGE

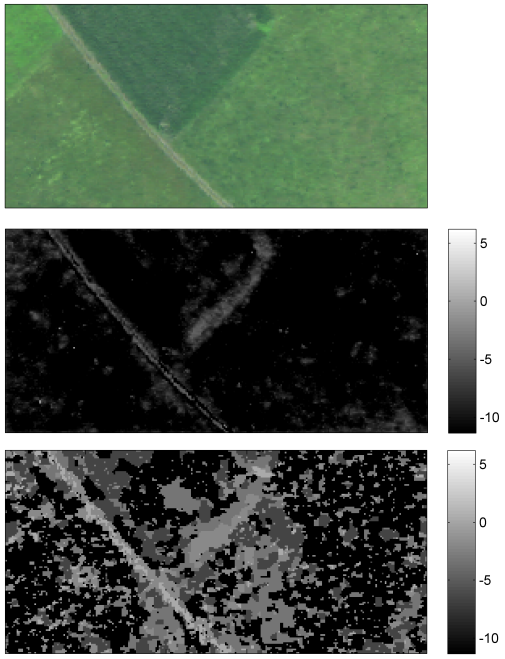

Fig. 1. Top: True color image of the scene of interest. Middle: Nonlinearity levels identified by G-RCA+ ($a = 0.18$). Bottom: Nonlinearity detection map obtained with RCA (5 nonlinearity levels).

This section presents an application of the proposed *Gamma-RCA* method with positivity constraints (G-RCA+) to a real hyperspectral image (we validated our method extensively using synthetic data and several real datasets, due to space limitations these simulations are available in the technical report [12]). The hyperspectral image considered in this experiment was acquired in 2010 by the Hyspex hyperspectral scanner over Villelongue, France ($00°03$'W and $42°57$'N). This scene was observed at $L = 160$ spectral bands ranging from the visible to near infrared with a spatial resolution of 0.5m. This dataset has already been studied in [13, 4] and is mainly composed of forested and urban areas (see [13] for more details about the data acquisition and pre-processing steps). We have applied our method to the region of interest of size 180×250 pixels that is depicted in Fig. 1 (top). This region is composed mainly of a path and different vegetation species and has $R = 5$ endmembers, whose spectral signatures have been extracted from the data using VCA [14]. We implemented G-RCA+ using $N_{\text{MC}} = 1000$, $N_{\text{bi}} = 500$, and by estimating α_3 by MML estimation (see the technical report [12] for more details). The computing time was 70 minutes on a 3GHz Intel Xeon quad-core workstation running MATLAB

2014, which is comparable to other MCMC-based Bayesian inference methods such as RCA [4].

Fig. 1 (middle) shows the nonlinearity levels estimated with the proposed G-RCA+ method. For comparison, the results obtained with RCA [4] are presented in Fig. 1 (bottom) (recall that to simplify the estimation problem, RCA artificially constrains nonlinearities to take a finite number of values). We observe that the results obtained with both methods are in good visual agreement and highlight spatial structures that can be easily identified in the colour image (e.g., path) where one would expect nonlinear mixing to occur. More importantly, by not constraining the number of nonlinearity levels, G-RCA+ produces spatially smooth estimates that appear realistic (and that do not require specification of the number of nonlinearity levels a priori). It is also important to note that the results obtained with G-RCA+ indicate that the nonlinear effects in the image are sparser and weaker than previously suggested by RCA. Indeed, in the model used in RCA, the nonlinearity coefficients $\mathbf{\Gamma}$ can take large positive and negative values which average out, leading to large estimated nonlinearity levels. By constraining the nonlinear coefficients to be non-negative, G-RCA+ yields smaller and sparser nonlinearity levels that are arguably closer to what could be expected.

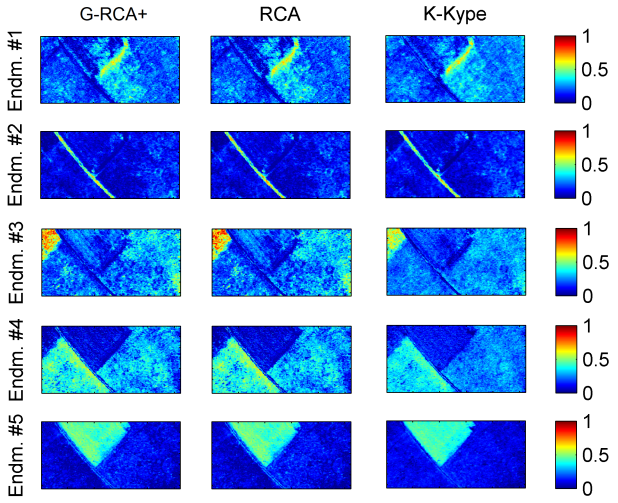

Fig. 2. Abundance maps estimated by G-RCA+ (left), RCA (middle) and K-Hype (right) for the Villelongue real image.

Moreover, Fig. 2 shows that the abundances estimates obtained with G-RCA+, RCA and with the state-of-the-art kernel unmixing algorithm K-Hype [15], that we implemented with a polynomial second order symmetric kernel with Gram matrix $\mathbf{K} = (\mathbf{M}\mathbf{M}^T) \odot (\mathbf{M}\mathbf{M}^T)$ and by optimising all the algorithm parameters by cross-validation. We observe that, despite being fully unsupervised, G-RCA+ produces abundance estimation results that are as good as the ones obtained with the state-of-the-art algorithms which are not fully unsupervised (the three methods achieved a reconstruction error of 0.22 [12]).

6. CONCLUSION

This paper presented a new hierarchical Bayesian algorithm for spectral unmixing of hyperspectral images which incorporates the spatial dependencies inherent in an image associated with the nonlinear mixture effects. The nonlinear mixtures were decomposed into a linear combination of the endmembers and an additive term which represents the nonlinear effects. This term was further decomposed as a combination of the endmembers cross-products. A Gamma Markov random field was introduced to promote smooth nonlinearity variations in the image. In contrast with previously reported work where nonlinear unmixing relied on a nonlinearity level-based image segmentation, the proposed model allows the level of nonlinearity to differ in each pixel while allowing the identification of regions where nonlinear effects occur. In this paper, a zero-mean Gaussian prior, restricted to the positive orthant was assigned to the nonlinear coefficients of each pixel. This choice was motivated by the fact that several existing models include positivity constraints for the nonlinear terms, e.g. [7, 8], include such constraints within the SU procedure, and this was previously not possible using the RCA model in [4] due to the marginalisation of these parameters. The results presented in this paper have shown that it can significantly improve the unmixing performance. In this paper, the endmembers were assumed to be perfectly known but often need to be extracted from the data. Future work will study generalizations of the G-RCA+ model that also account for imprecise endmembers, for the spatial correlations between their abundances coefficients, and for misspecified numbers of endmembers.

7. REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

[2] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 82–94, Jan. 2014.

[3] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 7, no. 6, pp. 1844–1868, June 2014.

[4] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Residual component analysis of hyperspectral images - application to joint nonlinear unmixing and nonlinearity detection," *IEEE Trans. Image Processing*, vol. 23, no. 5, pp. 2148–2158, May 2014.

[5] A. Kalaitzis and N. D. Lawrence, "Residual components analysis," in *Proc. Int. Conf. Mach. Learning (ICML)*, 2012.

[6] P. J. Green, "Reversible jump MCMC computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, Dec. 1995.

[7] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image and Signal Processing for Remote Sensing XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds. 2009, vol. 7477, p. 74770I, SPIE.

[8] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[9] O. Dikmen and AT. Cemgil, "Gamma markov random fields for audio source modeling," *IEEE Trans. Audio, Speech, Language Processing*, vol. 18, no. 3, pp. 589–601, March 2010.

[10] Marcelo Pereyra, Nick Whiteley, Christophe Andrieu, and Jean-Yves Tourneret, "Maximum marginal likelihood estimation of the granularity coefficient of a Potts-Markov random field within an mcmc algorithm," in *Proc. IEEE-SP Workshop Stat. and Signal Processing*, Gold Coast, Australia, July 2014.

[11] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, New York, second edition, 2004.

[12] Y. Altmann, M. Pereyra, and S. McLaughlin, "Bayesian nonlinear hyperspectral unmixing with spatial residual component analysis," *ArXiv e-prints*, Oct. 2015.

[13] D. Sheeren, M. Fauvel, S. Ladet, A. Jacquin, G. Bertoni, and A. Gibon, "Mapping ash tree colonization in an agricultural mountain landscape: Investigating the potential of hyperspectral imagery," in *Proc. IEEE Int. Conf. Geosci. and Remote Sensing (IGARSS)*, July 2011, pp. 3672–3675.

[14] José M.P. Nascimento and José M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, no. 4, pp. 898–910, April 2005.

[15] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, 2013.